# *Methodology*

*As a consequence of our membership in the European Union and to comply with EU legislation, namely the 1177/2003 framework Regulation and implementing regulations of the Commission, the Czech Republic conducts annual EU-SILC (European Union - Statistics on Income and Living Conditions) surveys on income and living conditions of households.*

*The 2020 round of the survey was carried out by the Czech Statistical Office (CZSO) in accordance with Act no. 89/1995 concerning the state statistics service and with Act no. 101/2000 on individual data protection. The official title of the survey was Living Conditions 2020.*

*The aim of the survey is to gather representative data on income distribution for the whole population and for various household types, data on housing - its quality and affordability, household durables, and labour, financial and health conditions of adults living in private households. Living Conditions 2019 contained a module focused on financial situation of households.*

## *1. Organization of the survey*

### *1.1 Sampling*

*The survey was carried out in all regions of the Czech Republic. The interviewers visited 6 622 dwellings whose questionnaires were completed in the previous year, 65 dwellings where a sample person from the previous wave moved to and 4 750 newly selected dwellings.*

*The sample was obtained by applying a two-stage probability sampling scheme to each of the 14 administrative regions (NUTS3 regions) independently. The total number of dwellings selected in each region was proportional to the region's size. At the first sampling stage small geographical areas (CEUs - census enumeration units or districts) were selected by probability sampling. These CEUs served as a basis for the second-stage selection (a sample of 10 dwellings was drawn from each CEU).*

*Before selecting the sample of dwellings, the sampling frame had to be adjusted to enable incorporation of small census enumeration units into the sampling process to reach the required full geographical coverage of the national territory. Small CEUs (with less than 20 inhabited dwellings) were merged with adjacent CEUs and the resulting larger CEUs entered the first stage of sampling. Therefore, in some cases, the 10 chosen dwellings could belong to two or more (in exceptional cases) CEUs.*

*The CZSO's regional fieldwork units (each covering one of the 14 NUTS3 administrative regions) received the list of selected dwellings (addresses + identification numbers of flats in apartment buildings). Before the actual fieldwork, the regional fieldwork units' staff carried out the identification of the selected dwellings and filled in the contact names on the list of selected dwellings for interviewers.*

### *1.2 Fieldwork*

*The sampling unit is a dwelling. During the first-wave visit all households and all the persons who have the dwelling as their main place of residence are surveyed. This rule also applies to foreign nationals and subtenants. During the waves 2-4 only those households are surveyed which include a panel person (those surveyed in the 1st wave). Households who moved from the original address are followed up. At their new address, all persons who are members of the same household as the panel person are surveyed.*

*Methods of acquiring data were carried out by dual scheme. A predominant part of the selected households was interviewed with an electronic questionnaire (CAPI); a part was interviewed using paper questionnaires (PAPI).*

*Data collection lasted from February 1 to July 26 2020. Collection of data was coordinated by workers from regional departments responsible for fieldwork. Workers from regional departments also conducted methodical training of fieldworkers.*

*The content of the survey was divided into four questionnaires with different units of reference:*

*Questionnaire A (dwelling unit questionnaire): contained the roster with the list of all persons with usual residence in the selected dwelling, their basic demographic characteristics, information on sharing of expenses to determine household units and relationship of each person to the main user of the dwelling and to the head of household.*

*Questionnaire B (household questionnaire): filled in for each household, contained information on housing, consumer durables, financial situation of the household, consumption of the household's own production (i.e. small scale farming and similar activities), inter-household transfers paid and received, family social benefits, rental income, paid regular taxes on wealth (buildings and land) and childcare.*

*Questionnaire C (personal questionnaire): filled in by each household member aged 16+ as of 31 December 2019 (i.e. persons born in 2003 and earlier). This questionnaire contained information on labour status and employment, personal income (from employment, private enterprise and social security schemes), participation in private pension plans, questions relate to the informations about material deprivation and health.*

## *1.3 Processing of the questionnaires and collected data*

*Regional survey coordinators were responsible for collecting the PAPI questionnaires from interviewers and also for entering the data into electronic CAPI questionnaire (including additional coding of certain questions, e.g. occupation (ISCO), economic sector (NACE), country of birth and citizenship.*

*The regional coordinators than merged all the CAPI data from interviewers and after some initial checking they sent the result to the CZSO central database for further processing. All data was handed over by the regions to the headquarters in an electronic form.*

*Methodologists at CZSO headquarters then perform final super-controls and central processing on complete data sets with questionnaire data and produce clean data files.*

## *1.4 Successfully interviewed households and non-response*

*The fieldwork revealed that among the total of 11 437 dwellings in the sample there were 701 (6,1 %) dwellings either unoccupied, or the address did not exist or the survey was not possible, e.g. because the households had moved. Since substitution for the ineligible units is not allowed, the survey was conducted in 10 736 dwellings and 10 769 households (there was more than one household in some of the dwellings).*

*The overview of the survey response is presented in the following table:*

***Tab. 1 Successfully interviwed households and non-response***

| | Households | | | Response rate (%) | | |
|---|---|---|---|---|---|---|
| | Total | 1st wave | 2nd-4th wave | Total | 1st wave | 2nd-4th wave |
| **Response, total** | 8 618 | 2 370 | 6 248 | 80,0 | 55,7 | 95,9 |
| **Non-response, total** | 2 151 | 1 885 | 266 | 20,0 | 44,3 | 4,1 |
| *refusals (unwillingness to give information)* | 1 599 | 1 413 | 186 | 74,3 | 75,0 | 69,9 |
| *household not contacted, temporarily absent* | 327 | 291 | 36 | 15,2 | 15,4 | 13,5 |
| *household unable to respond (health limitation)* | 98 | 80 | 18 | 4,6 | 4,2 | 6,8 |
| *other reasons (linguistic etc.)* | 127 | 101 | 26 | 5,9 | 5,4 | 9,8 |

*Source: CZSO*

*Refusals also include situations when the household did not refuse the survey as such, but did not agree to provide the information on income to an extent that would qualify the household as successfully interviewed. The definition of successfully interviewed household allowed missing income data for only one person that is not the household's head.*

*The category comprising non-contacts or those temporarily absent covers situations when the interviewer was not able to contact the selected household, despite having made the prescribed minimum number of attempts at personal contact.*

*The overview of the survey response in the CR and its regions is showed in the following table:*

***Tab. 2 Survey response in the CR and its regions***

| Region (NUTS3) | Total | | | 1st wave | | | 2nd-4th wave | | |
|---|---|---|---|---|---|---|---|---|---|
| | HHs in survey | response | | HHs in survey | response | | HHs in survey | response | |
| | | count | % | | count | % | | count | % |
| *Hl. m. Praha* | 1 351 | 926 | 68,5 | 612 | 238 | 38,9 | 739 | 688 | 93,1 |
| *Středočeský* | 1 232 | 989 | 80,3 | 492 | 269 | 54,7 | 740 | 720 | 97,3 |
| *Jihočeský* | 684 | 596 | 87,1 | 249 | 170 | 68,3 | 435 | 426 | 97,9 |
| *Plzeňský* | 599 | 475 | 79,3 | 235 | 125 | 53,2 | 364 | 350 | 96,2 |
| *Karlovarský* | 273 | 209 | 76,6 | 121 | 60 | 49,6 | 152 | 149 | 98,0 |
| *Ústecký* | 784 | 634 | 80,9 | 322 | 193 | 59,9 | 462 | 441 | 95,5 |
| *Liberecký* | 422 | 365 | 86,5 | 156 | 106 | 67,9 | 266 | 259 | 97,4 |
| *Královéhradecký* | 529 | 432 | 81,7 | 201 | 118 | 58,7 | 328 | 314 | 95,7 |
| *Pardubický* | 555 | 459 | 82,7 | 207 | 131 | 63,3 | 348 | 328 | 94,3 |
| *Vysočina* | 564 | 466 | 82,6 | 200 | 118 | 59,0 | 364 | 348 | 95,6 |
| *Jihomoravský* | 1 257 | 1 011 | 80,4 | 487 | 279 | 57,3 | 770 | 732 | 95,1 |
| *Olomoucký* | 601 | 477 | 79,4 | 238 | 129 | 54,2 | 363 | 348 | 95,9 |
| *Zlínský* | 584 | 474 | 81,2 | 223 | 128 | 57,4 | 361 | 346 | 95,8 |
| *Moravskoslezský* | 1 334 | 1 105 | 82,8 | 512 | 306 | 59,8 | 822 | 799 | 97,2 |
| *CZ total* | 10 769 | 8 618 | 80,0 | 4 255 | 2 370 | 55,7 | 6 514 | 6 248 | 95,9 |

*Source: CZSO*

*Participation in the sample survey is voluntary; unlike the population census, households were not obliged to provide any information. A selected household has to be informed about the content of the survey and about the fact that their participation in the survey is voluntary. Whether to respond or not is left to the household's own deliberation. The main reasons for refusal are privacy reasons (objections against giving personal information and fear of abuse of personal data), fear of contact with interviewers as strangers. There is a considerable group of persons, who as a matter of principle strictly refuse to give any information.*

## *1.5 Grossing up and weighting*

*When compared with the data from other statistics and registers, selected characteristics of our sample showed that a phenomenon typical of household surveys had occurred - high level of non-response (in a rotational panel influenced by a prior response) had biased the proportions in the final data file from which results are obtained. The deformation of demographic characteristics and social structure of the sample did not allow us to use simple techniques of grossing up (post-stratification). To reach a sufficient level of bias elimination, which is the necessary pre-condition for obtaining good estimates, it was necessary to use more sophisticated methods.*

*In practice, the well-tried iteration method of weight calibration was utilized, which minimizes the difference between the known and the grossed up values of selected characteristics. Although it is a panel survey comprising data of four practically independent samples (waves 1-4), a simple calibration method was utilized which did not distinguish the waves but worked with all households together.*

*At the same time and according to the Eurostat's recommendations the standard system of integrated weights was used in the survey, i.e. a single set of grossing-up coefficients that was subsequently used to produce results for both households and individuals.*

*As the basis for calculations the following calibration variables were used:*

- *Number of inhabited dwellings in each NUTS3 region, subdivided into family houses (detached and semi-detached houses) and apartments, based on the 2011 Census continuously updated from administrative sources of construction authorities*
- *Population characteristics:*
  - *Population totals in each NUTS3 region (from demographic statistics)*
  - *Economic activity characteristics in each NUTS3 region*
    - *Number of employees – derived from the number of employees in the economy based on the Labour Force Survey (LFS) results and company reporting*
  - *Economic activity characteristics in each NUTS3 region:*
    - *Number of pensioners (excl. pensions for orphans) - based on the administrative data from the Ministry of Labour, Social Affairs and the Czech Social Security Administration and reduced the pensioners living out of the dwellings based on the 2011 Census*
    - *Number of unemployed - registered unemployment from the administrative source of the Ministry of Labour and Social Affairs, corrected for unregistered unemployment using the Labour Force Survey data and for unemployment of the homeless and persons living in institutions or collective accommodation establishments (based on the 2011 Census)*
    - *Number of recipients of parental allowance – data from Ministry of Labour and Social Affairs for 2018*
    - *Number of self-employed - estimate based on the Labour Force Survey and on the administrative data from the Czech Social Security Administration*
    - *Number of children aged 0-15 - from demographic statistics*
  - *Demographic characteristics in each NUTS3 region (based on the demographic statistics):*
    - *Age groups (0-15, 16-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+); Sex*
  - *Municipality size (less than 2 000 inhabitants, 2 000-9 999, 10 000-49 999, 50 000 or more inhabitants)*

*The target population of the survey was persons living in private households, therefore the data from demographic statistics was adjusted by subtracting institutionalized population (from social security administrative data and Ministry of Justice) and the persons living outside dwellings as based on the 2011 Census.*

*As the sampling unit is the dwelling, all weight coefficients were calculated for dwellings and subsequently assigned to all persons and households in them (integrated weights).*

*The method described above deals with non-response successfully, i.e. it corrects the bias due to the specific composition of households that did not respond. First of all, it improves demographic and social structure but, as a by-product, it also eliminates deformation of income indicators related to these structures.*

*Another source of bias, which needs to be taken into account, stems from the method of interviewing. Data on income and housing costs obtained during face-to-face interviews with household members has the tendency to underestimate or overestimated, certain income sources or data on some income components can be completely missing (item non-response). Not to reduce the size of the processed dataset the missing income was imputed using correct statistical methods.*

*In Living Conditions 2020 the interviewer failed to obtain income information for one person in an otherwise successfully interviewed household only in 14 cases. The missing income of such individuals was replaced with income of another randomly selected person with the same characteristics, i.e. a simple hot-deck method was applied.*

*Underestimation of income is a natural consequence of the fact that respondents either tend to state lower than actual values or simply do not recall having had certain irregular or small incomes at all. It is, more or less, a non-sampling error, affected substantially by the incomes themselves and by their source. The possibilities to eliminate this underestimation are limited. In the presented survey, only such adjustments were made where there was a sufficiently reliable external statistical source or where the conjectures could be based on legislation.*

*If the respondent reports income from employment as net, the net income often shows a significant tendency to be distorted (either under- or overestimated) and the non-sampling error grows. This might occur*

*when the employer deducts a certain amount from the employee's wage/salary (e.g. alimonies or pension scheme contributions). When calculating the gross income, this bias is usually adequately compensated for by using additional information from the survey. The level of gross income from employment was compared with the results of wage statistics and for persons who were revisited and stayed in their former jobs with data from the previous waves. Some respondents mistakenly reported gross income as net or vice versa and thus there were significant and inexplicable year-to-year differences. In such cases top-bottom coding was applied or the data were edited. With the self-employed no income corrections were necessary.*

*In the case of social benefits for which there is legal entitlement (parental leave, child birth benefit, death grant provided to families of the deceased, to some extent also maternity leave), a check on their receipt by eligible households was applied and amounts provided were corrected according to the amounts set by the legislation. With old age benefits (pensions from the social security system) the tendency to underestimation is negligible but as there were falls or increases in this kind of income without any outward reason, the amounts were corrected according to the last year's values.*

*It was not possible to correct the underestimation of sickness benefits (omissions related to short-term illnesses could not be identified in the existing data), means-tested social benefits whose claims depend on the previous income (prior to the income reference period), capital income, or income in kind and inter-household transfers.*

*The comparison of the aggregated income from this survey with the household sector aggregates of the national accounts (even after subtraction of items which are not covered by household income surveys) is problematic. Concerning its aggregated value the income obtained by direct questioning in households will always be lower. More important fact for evaluation of their credibility is that the trend in the development of household income follows trends in the national accounts. From this viewpoint, the presented results of Living Conditions 2020 are reliable and, as to their time series, consistent. They are fully comparable with similar statistics produced in the EU states.*

# *2. Methodological notes on published tables*

## *2.1 Basic definitions*

*The publication contains the results for households and individuals aged 16 and older. The definition of household is based on the sharing of expenditures concept, in line with the definition of Paragraph 115 of the Civil Code - based on the declaration of the persons in the dwelling that they permanently live together and pool their finances to cover their needs. As the 16 year olds those persons were regarded who had reached this age by 31 December 2019.*

*Reference periods:*

- *Demographic variables - age: 31 December 2019; marital status, education, housing, financial situation: the date of the interview.*
- *Work activity was collected for each month of 2019 as well as currently. Work activity figures are gathered by self-definition of the respondent (respondents themselves choose among different types of activity the one that fits them the most). Its value primarily depends on the respondent's main occupation and on the time spent in it. Subsequently, other data related to the respondent's work activity (status at work, profession) were collected as of the date of the interview. Parallel activities were surveyed (second job, study).*
- *Economic activity was not collected but derived from the monthly/yearly data (if monthly data was the basis, the activity with the highest incidence was coded as the yearly value). For those who completed their education in 2019 the latter half of the year was considered.*
- *Income data (both monetary and in kind): calendar year 2019.*
- *Subjective questions focused on housing and financial problems: the date of the interview.*
- *Health problems: last six or twelve months.*
- *Housing, consumer durables, financial and social situation of household: the date of the interview, unless the question specifically refers to some other period.*

## *2.2 Description of variables*

### *2.2.1 Household composition*

- *Size of the household - the number of household members on the date of the interview, including persons temporarily away, if the period of actual or foreseen absence is shorter than 6 months and the person has no other private address. For persons studying away from home, the period of absence may be longer than 6 months, provided that the person has no private address and retains financial ties to other household members. Persons with a period of absence longer than 6 months, persons without financial ties to the household and persons temporarily present at the time of the interview who have their private address elsewhere are excluded.*
- *At work - during 2019 the prevailing economic activity status of these persons was employed (employees, self-employed, members of production cooperatives, unpaid family workers in family businesses). Persons drawing sickness benefits, students who apart from their study also worked (in employment, private enterprise), pensioners or persons on maternity leave with regular income from work were also included.*
- *A subgroup of self-employed persons was taken out of persons at work and defined as a separate group. The group consists of persons who do business based on trade licenses or special regulations, participants in joint ventures based on contracts (Limited liability companies, Corporations), professionals (doctors, lawyers, tax advisers) and persons working for royalties (artists, interpreters).*
- *Dependent children - children under 26 years of age, provided that the person is still in education despite having their own income (orphan's pension, education-related allowances, social benefits, income from short-term work); furthermore, those children were also included who due to a mental or physical handicap had not been able to attend school or earn their own living and who had not yet received their invalidity pension.*
- *Pensioners (without economic activity) - persons receiving pensions from the social security system (old-age, disability, survivor's) or persons of retirement age that were not eligible to receive pensions and they had not a job at all or their working activity had only limited extent.*
- *Unemployed - persons who did not have a job for a predominant part of the year but who wished to have one. Such persons did not have to comply with the strict ILO definition about actively seeking a job and readiness to start one.*
- *Persons on parental leave - persons who were employed before parental leave and were guaranteed a return to that job during parental leave.*
- *Other persons - inactive persons caring for household members in need of care, persons living on property income and others.*
- *Incomes are presented as household incomes, per capita incomes or equivalised incomes - using OECD and modified OECD equivalence scale. Household income is equivalised in order to reflect differences in household size and composition. The equivalised income is defined as a per-equivalent-adult figure. In other words, to obtain a person's equivalised income, the total household income is divided by its equivalent size. The OECD equivalence scale gives a weight of 1,0 to the first adult, 0,7 to any other household member aged 14 and over and 0,5 to each child below 14 years. The modified OECD equivalence scale gives a weight of 1,0 to the first adult, 0,5 to any other household member aged 14 and over and 0,3 to each child below 14 years.*

### *2.2.2 Household characteristics*

- *Head of household - for couples with or without children it is always the male, regardless of his economic activity. In lone-parent families (one parent with child/children) and in non-family households (persons not related by marriage or partnership, or parent - child relationship) the first criterion for determining the head of household was economic activity and the second criterion was income of household members. This rule was also applied in more complicated household types (for example in the case of sharing expenditures among multiple two-parent families).*
- *Household type - is based on household composition. Two-parent families are based on a couple (married or cohabitating), with or without children. The lone-parent family category contains households with one parent and at least one child. In addition to these basic structures, these households may also contain other household members. The households where all children are dependent and there are no other members except parents are labelled as nuclear families. One- or two-parent non-nuclear families*

*comprise also other persons than just parents and their dependent children. Non-family households are entities consisting of two or more persons with other than husband-wife or parent-child relations.*

- *EU type of household - in contrast to the previous definition, this typology does not depend on family relations and is only based on number of adults and dependent children. Dependent children are all persons 0-17, and, further, persons 18-24 who are economically inactive, are not looking for a job or available for work and live with at least one parent. Households of individuals and two adults were further divided into age groups: one person household aged less than 65 years, one person household aged 65 years or more, two adults both under 65 years and two adults with at least one person aged 65 years or more.*
- *Education - 4 categories: primary (incl. incomplete), lower or vocational secondary, upper secondary, tertiary. Upper secondary education includes also vocational education with a secondary school-leaving exam. Post-secondary non-university education falls under the rubric of tertiary education. Besides this, tertiary education includes all tertiary programs - bachelor, graduate and post-graduate level.*
- *Occupation - 9 main classes of national classification CZ-ISCO. Households are classified according to the occupation of the head of household. Soldiers were coded as 1 - managers and legislators.*
- *Household group was based on the status of the head of household.*
  - ***Households, total** - represent the average household in the Czech Republic,*
  - ***Households of employees** - the household head's prevailing activity status is employee*
    - *Households of employees with lower education: the highest attained education level of the head of household is primary or secondary-vocational,*
    - *Households of employees with upper education: the highest attained education level of the head of household is complete secondary or tertiary,*
  - ***Households of self-employed** - the household head's prevailing activity status is self-employed (in whatever field it may be, including agriculture),*
  - ***Households of pensioners** - the household head was an inactive pensioner; this group is further divided into two subgroups based on whether there is anybody in the household who worked,*
  - ***Households of unemployed** - household head's prevailing activity status is unemployed (at the same time, in two-parent families the female partner or grown-up children can by employed),*
  - ***Other households** - household head's prevailing activity status is other than one of the four previous categories (for example a person on parental leave benefit, students or persons living on property).*
- *Subsistence minimum is based on the amounts of national subsistence minimum applicable in 2019. The monthly subsistence minimum is the sum of amounts pertaining to the individual household members.*

***Tab. 3 Subsistence minimum, 2019***

| ***Subsistence minimum (CZK/month)*** | | | |
|---|---|---|---|
| *One person household* | *3 410* | *Dependent children:* | |
| *Multiple-member households:* | | *-6* | *1 740* |
| *first person in household* | *3 140* | *6-15* | *2 140* |
| *other adult persons* | *2 830* | *15-26* | *2 450* |

***Source: MoLSA***

### *2.2.3 Monetary and non-monetary income*

*Incomes related to household as a whole were collected at the household level. There were social benefits targeted at households, rental income and value of goods produced directly by the household through either private or professional activities.*

*Incomes collected at individual level: income from employment (main job, secondary jobs) incl. other income related to them (remunerations, shares, bonuses), income from contracts, income from self-employment, sickness benefits, old-age benefits, unemployment benefits, social benefits attributable at individual level (such as parental benefit or disability benefits) and other incomes (capital income, insurance claims).*

*Income from employment (both main job and possible secondary jobs) was collected both either gross of tax and social insurance or net, incomes from contracts only gross. Self-employed persons could choose from several ways to record the result of their enterprise. They could state the gross profit/loss according to their tax declaration, they could give the sum which served as the yearly basis for calculating their monthly health and social security contributions or could make their own estimate of their gross or net*

*profit/loss. Family members co-operating in private enterprise run by another member of the family stated only proportionate part of the income from the business.*

*Rental income was collected either gross or net, based on what information respondents were able to provide. All other kinds of income were collected net and subsequently appropriate rules of the tax system were applied to estimate the gross amounts. In addition, the information was collected on claimed tax deductibles to enable calculation of taxes and social insurance contributions. Sum of individual net incomes then forms the main national indicator – net money income of the household.*

*Besides this national indicator of household income, it was necessary to construct an internationally comparable household income indicator, which is based on Eurostat methodology for EU-SILC surveys. This indicator is called disposable household income. The difference between these two definitions of the household income is in inclusion/exclusion of certain components of income ( lump sum and irregular inter-household transfers, non-cash employment income, regular taxes on wealth).*

*Household income in kind consists of consumption of food, products and services originating from the household's own production activity (for example food or domestic animals from own farm, value of food from own restaurant, value of timber from own forest) and of perquisites provided by employers (company car, company-paid or co-financed meals and other non-cash paid services). The CZK value of own-production in kinds was calculated from reported amounts using the average price of the given commodity. The amount of CZK 3000 was added to income in kind of an employee for each month of using a company car. The financial contribution of the employer to the employee's meals was calculated from the annual or monthly estimate of the respondent. Because these contributions form an important part of employees' income, they are tabulated under a separate heading (Tab. 1a).*

*Detailed income components are presented in Tab. 1. The breakdowns in other tables are less detailed. More detailed breakdowns are provided for gross income.*

*Selected income components:*

- *Income from employment was defined in line with the national tax law. It includes income from employment based on a contract or similar arrangement between employer and employee. It also includes incomes of owners of the incorporated business from work for their company, income of members of statutory boards and other governing bodies of corporations, remuneration based on holding of elected public posts, income of apprentices in vocational schooling for their work undertaken as a part of their practical training and income from flexible short-term contracts under special regime set in the Labour Code. Using company car for private purposes is also classified as income in kind from employment.*
- *Income from self-employment includes also income from farming activities, if these are conducted as a business activity, income from independent professional practices (lawyers, doctors) and income from intangible assets (copyrights, royalties).*

*Income from main employment includes income of employees from their main job. For multiple coincident jobs, the declaration of the main job was left to the respondent.*

*Income from secondary employment includes salaries from secondary jobs, conducted besides the main job or self-employment activity of the respondent and income from flexible short-term contracts under special regime set in the Labour Code.*

*Income from secondary self-employment activity is analogous to the secondary employment income. It includes income from secondary self-employment activity undertaken in addition to the main job of the respondent (where respondent declared employment contract as his/her main job).*

- *Social income is in principle net. Gross amounts were included for cases of pensions above the tax-exempt limit. In these cases, tax was applied to the amount above this limit (CZK 480 600).*
- *Sickness benefits include all sorts of benefits from the social sickness insurance, i.e. maternity leave benefit, reduced employment income compensation in pregnancy and motherhood, income support for persons caring for household member in need of short-term care (mostly care for children during their illness). Sickness benefits include work inability compensation paid by the employer.*

*Other social support benefits include social benefits for foster parents taking care of adopted children, birth and death grants.*

*Other social benefits include certain benefits connected to the termination of employment in selected professions, various other social benefits like benefit for persons providing long-term homecare for their*

*relative in need, support for care in spas and other social benefits for families with children, old and disabled citizens, which are mostly administered by the municipal authorities.*

*Social exclusion allowances include regular and lump sum monetary benefits that help the household pay their food and housing bills, or contribute to satisfy their basic needs.*

*Scholarships include all kinds of scholarship money income from schools and, furthermore, pocket money paid to apprentices by schools or future employers.*

*Social income from abroad although the benefactor is not the government of the Czech Republic went under respective rubrics and was mixed with the Czech government's help (pensions and child benefits).*

- *Other income*

*Income from capital contains interest from savings, bonds and various forms of deposits, dividends from shares, profits from incorporated businesses, income from investments abroad.*

*Other income includes income from occasional property rentals, life and material insurance, sale of own-produced goods, income from organisations not elsewhere classified (scholarships and pocket money of apprentices, grants from charity and non-governmental organisations), lottery winnings, prizes, pay for occasional not contracted jobs, regular inter-household transfers (alimonies and the like).*

#### *2.2.4 Housing costs*

*In the case of more than one household in one dwelling unit, the costs were divided according to their actual contribution to their financing.*

*When the household reported its housing costs only in one item as the rent paid for accommodation, the partial amounts were estimated based on the data from households, which provided the detailed information on their housing costs. Estimates were modelled by regression models taking into account the type of dwelling (family houses vs. other), type of rent (market rent vs. reduce rent contracts), number of household members and usual local level of housing costs (municipality, census enumeration unit).*

## *2.3 Data tables - description and notes*

*The publication contains tables for households (Tab. 1 to 15), for persons aged 16+ (Tab. 16 to 18) and tables of poverty and material and social deprivation (Tab. 19 and 20).*

*Values in the tables were calculated from the weighted microdata and rounded. The total counts of households or persons may therefore not always exactly correspond to the sum of the counts for a given breakdown. For the same reason, the sum of percentages may not always be equal to 100.*

*Whenever the term "children" is used in table headings, it always means dependent children (in accordance with the definition in part 2.2.1), except tables 9, 19 and 20 where the EU definition of dependent children is used. Net household income was used for all income-based classifications.*

*The publication tables with data on households are designed to correspond to the tables from previous years of the survey. This enables the user to compare results over a longer span of time. The legends of the tables show clearly what population (or subpopulations) of households the results are calculated for. The set of tables is divided into five parts labelled by name and letter a) to e), each part with its own legend:*

a) *Basic data on household composition and income. Income data are mainly per capita averages; in selected tables the presented income data are equivalised using OECD and modified OECD equivalence scales. Presented averages of equivalents enable users to calculate equivalised income based estimates also for the rest of the breakdowns. Table 1 contains a more detailed breakdown of income; other tables are restricted to only main income components. The variable "tax bonus" represents amounts of tax help provided to low-income families with children. Due to the tax bonus net income of such families is in most cases higher than the gross one. The relation between net and gross income can be expressed by the following formula: net income = gross income - insurance - tax + tax bonus.*

b) *Income distribution of households and persons for fixed income groups. The income group's brackets are adjusted to reflect net monthly per capita income distribution or monthly disposable per modified OECD scale unit income distribution, respectively. Incomes are further related to the minimum subsistence level and to the median of per capita income for the whole population. The b) part is absent in those tables where income per capita was used as the classification criterion.*

c) *Characteristics of households, which describe their structure according to the various classification criteria and which supplement or explain data on income.*

d) *Characteristics of housing of given household groups, equipment with selected consumer durables and housing costs, which are presented as monthly averages per household. In 2017, a question on prevailing heating was added to the table. The household equipment was newly enlarged with a clothes dryer and dishwasher, on the contrary the colour TV, washing machine and phone was removed.*

e) *Opinions of households on their housing, financial problems - for example in connection with the housing costs, repayment of loans and ability to make ends meet*

### *2.3.1 Notes on selected tables with household data*

*Tab. 1 gives data for household groups, which are comparable in long time series. It offers a look at changes in household structures, their demographic characteristics and incomes.*

*Tab. 2 to 4 - households total by deciles based on net monetary income per capita and modified OECD equivalence scale, households of employees and households of pensioners by quintiles based on net monetary income per capita. The households were ordered and divided into deciles/quintiles according to the net per capita household income, or net equivalised household income using the modified OECD equivalence scale. The values of deciles and quintiles are incomes of the last household in that quintile group. While grossing up the survey data, it is not possible to maintain exactly the same number of households in each quintile group. Therefore, the household counts may slightly differ.*

*Tab. 5 is the result of comparison of the monthly net household incomes after deduction of housing costs with their subsistence minimum from the national law on social need (as of 2019). The multiples of the subsistence minimum were adjusted to correspond with social security benefit entitlement limits.*

*Tab. 6 and 7 comprise information on households broken down by number of dependent children and number of household members at work.*

*Tab. 8 presents a breakdown of childless households by at-work status of their members.*

*Tab. 9 - using EU-compatible typology, the classification enables international comparisons.*

*Tab. 11 - a size of municipality, as of December 31, 2019, from demographic statistics.*

*Tab. 13 - a type of household and the highest education attained by its members. Only households where the head of household is economically active were included. In two-parent families, the education of the head of household is combined with the education of his spouse. Some low frequency combinations are omitted. Primary education includes secondary-vocational education and persons with incomplete primary education.*

*Tab. 15 - time series of selected household indicators since 2015. The income reference period is the previous calendar year.*

### *2.3.2 Notes on 16+ tables*

*Tab. 16 to 18 - data are classified according to the demographic characteristics and the size of the municipality where the persons live. In addition to the presented basic economic activity variables, the prevailing part of the table presents data on subjective evaluation of health. This part does not include proxy respondents (respondents, for whom the questionnaire data was collected from another household member) - since proxy answers were not allowed for this part of the personal questionnaire.*

*Tab. 19 - contains time series of person's at-risk-of-poverty rate since 2015.*

*The calculation of the at-risk-of poverty rate is based on the equivalised disposable income, which is a ratio obtained by dividing the disposable income of a household by the number of its adult equivalents (modified OECD scale). The resulting equivalised income is assigned to all the members of the given household (all persons in the household have the same income). The poverty threshold is then identified in the dataset of all persons ordered by their equivalised income. The most frequently used threshold is defined as 60 per cent of the median equivalised income. There are time series of the at-risk-of poverty rate using 60 per cent and 70 per cent threshold in the table 19. The at-risk-of-poverty rate is then the percentage share of people with equivalised income lower than the threshold in the population group we are interested in (e.g. males or old-age pensioners). This methodology is applied throughout the EU states and enables international comparisons within the EU.*

*Prevailing economic activity in tables 19 and 20 corresponds to the definition already mentioned in chapter 2.1. For inclusion in this part of the table, the activity has to last at least 7 months of the reference year 2018. Persons not fulfilling this condition were not included in this calculation.*

*Tab. 19 is supplemented by selected indicators of income poverty, which characterize variability of income in more detail.*

*Quintile share ratio S80/S20 - is the ratio calculated as the sum of equivalised income of the 20 percent of percent with the highest income (fifth quintile) divided by the sum of equivalised income of the 20 percent of persons with the lowest income (first quintile). Higher values of this ratio mean higher differentiation of incomes.*

*Relative median at-risk-of poverty gap - is the difference between the median income of those under the poverty line and the value of the poverty threshold expressed as a percentage of the poverty threshold. Higher value of this indicator means a deeper fall of persons under the poverty threshold. This indicator was calculated for the poverty threshold at 60 percent of the median.*

*Gini coefficient - is calculated from the ordered distribution of equivalised income. It reflects the relationship between the cumulative proportions of persons and cumulative proportions of income. Its values are in the range of 0 to 1. The higher the value of the Gini coefficient is, the higher differences in income of persons are. In publications it is usually presented in percents.*

*Tab. 20 presents data on material and social and material deprivationof persons broken down by some personal and household characteristics. According to the Eurostat's methodology, materially deprived persons are those who reported deprivation in 4+ out of 9 stipulated items. Questions concerning material deprivation are focused on consumer durables and financial limitations and difficulties.*

*Material and social deprivation rate is an indicator deriving from material deprivation which additionally takes into account the social situation of the household members. It is based on 13 items in total. As materially and socially deprived a person is regarded who could not afford at least five of the 13 items due to financial reasons. The person-related items include personal articles (possession of two pairs of shoes, new clothing), eye-to-eye contact with friends and family, leisure-time activities, the possibility of personal consumption expenditure and Internet connection.*

*The table shows results pertaining to the total population, to those at risk of income poverty and to those not at risk of income poverty.*

## *3. Results accuracy*

*When interpreting and analysing the results of the Living Conditions survey, one has to keep in mind the fact that the results are based on sample survey data and only subsequently inferred to the whole population. It means that all published data are statistical estimates based on a survey sample and comprise possible sampling and non-sampling errors.*

*The non-sampling error occurs in all surveys and censuses as well. It might occur as a consequence of many reasons, mostly of inaccurate methodological instructions, not respecting them by interviewers, wrong wording of questions, processing mistakes, unwillingness to participate in the survey or giving purposely biased answers. By meticulous care in all phases of data collection and processing one can reduce this component of total bias significantly. However, it is difficult or nearly impossible to evaluate its influence on the results. With the precondition of well-defined auxiliary variables, one can compare their distributions in the sample with the known distribution in the whole population (census).*

*The sampling error is a consequence of processing the results of not all units of the population, but of the sample data only. One has to infer the figures for the whole population from the obtained survey results. It can be evaluated using the sampling survey theory. This error can be limited by choosing a sample which is large enough and representative. Other factors can also influence the sampling error, namely sampling design, incidence of the measured variable and its natural variance.*

*Relatively low willingness of households to participate in the survey has been a persistent problem. In the case of repeated visits in the panel it results in a narrower range of types of households in the data collected and processed. This bias is corrected by calibration techniques described in the chapter 1.5.*

## 3.1 Estimates of sampling errors, confidence intervals

*There are two ways to evaluate the sampling error: either by a point-estimate of variance or by a confidence interval for the observed variable. Mostly 95% confidence intervals are constructed with a radius that we get by multiplying the standard error by a quantile of normal distribution - 1,96. The theory says that with the probability of 95% the actual value of the measured variable will lie within this interval. In this publication the measured variables are either frequencies - both relative and absolute - of how many households show a certain characteristic or means/totals of incomes.*

*Sample survey theory distinguishes between two types of totals - population totals and sub-sample totals. The sub-samples are the results of applying various criteria to the whole population like dividing the whole population into specific household groups according to the head of household's economic activity.*

*When ascertaining the sampling error, the biggest problem is the standard error calculation, which is done for each estimate separately. Computing standard errors of percentage totals or relative frequency is the easiest. A relative frequency can mean e.g. number of households of self-employed members as a percentage of all households. In the case of other estimates (e.g. income totals and their means per household or per capita) one must compute the standard error directly from the primary data and for each sub-sample separately. The tables illustrate the volatility of variability of various indicators, different sub-samples and several types of income.*

## 3.2 Confidence intervals for frequencies

*The following two formulas are simplified approximations of exact formulas and are applicable only to variables with binomial distribution. They thus apply to incidence estimates, like the percentage of incomplete families. In such cases, the deviations between the approximations and exact formulas are statistically insignificant. However, the formula for sub-population totals (onward characteristic A) might give inexact results for small area estimates; therefore the values in the upper left corner of the Tab. II were omitted.*

*Both formulas can be used for computation of confidence interval of random variables with binomial distribution:*

*a) for the <u>population total</u>*

*95% confidence interval of estimate $Y_A = y_A \mp 1{,}96.\ s_{yA}$, where*

$$s_{yA} \doteq N \cdot \sqrt{(1-f).\frac{\frac{y_A}{N}.(1-\frac{y_A}{N})}{f.N}} \qquad (1a)$$

*N* is the population size,
*f* is the relative sample size (n / N),
$y_A$ is the estimate of total $Y_A$ of characteristic A in the population

*Note: In the case of estimating confidence interval of relative frequency, one should substitute the relative frequency for the ratio* $\frac{y_A}{N}$ *in the numerator.*

*b) for the <u>sub-population total</u> (of observed characteristic B on the set of A)*

*95% confidence interval of estimate $Y_{AB} = y_{AB} \mp 1{,}96.\ s_{yAB}$, where*

$$s_{yAB} \doteq y_A \cdot \sqrt{(1-f).\frac{\frac{y_{AB}}{y_A}.(1-\frac{y_{AB}}{y_A})}{f.y_A}} \qquad (1b)$$

$y_A$ is the estimate of total $Y_A$ of characteristic A in the population,
*f* is the relative sample size (n / N),
$y_{AB}$ is the estimate of total $Y_{AB}$ of characteristic B on the set of A.

*Note: Again, one can substitute the respective relative frequency of characteristic B on the set of A for the ratio* $\frac{y_{AB}}{y_A}$ *in the numerator.*

***How to use the attached tables for determining frequency confidence intervals***

***Tab. I*** *Estimates of 95% confidence intervals of population totals for households and persons in the Czech Republic*

*The table is designed to determine an approximate 95% confidence interval of population totals of frequencies from the set of households or from the set of persons at the level of the Czech Republic as a whole. Let us take an example. In "Tab. 1 – Households by the status of head of household" we can find an estimate of 854,6 thousand households of the employees with lower education, and want to know the confidence of this estimate. In Tab. I we look up (in column "Households - estimate - thousands") the row closest to this number, namely 900. In this row we find the particular confidence interval, which in this case amounts to ± 37,78 thousand, for relative frequency the confidence interval is 20,16 ± 0,85 %. Because the number of households falls almost in the middle of the 800 - 900 interval, it is suitable to make the value more precise by using simple linear interpolation. Then the confidence interval expressed in absolute terms is 850 ± 36,95 thousand (average of 36,11 and 37,78) and, relatively, 19,04 ± 0,83 %.*

***Tab. II*** *Estimates of 95% confidence intervals of subpopulation totals for households*

*The table determines an approximate 95% confidence interval of subpopulation totals of frequencies from the set of households at the level of the Czech Republic as a whole. So provided we want to find out the confidence of the estimate of frequency of lone-parent families in households of the self-employed, which was 7,9 % of 575,3 thousand, we will look up in Tab. II the row closest to the number 575,3 namely 600 and the column closest to the number 7,9, namely 8. The confidence interval for the relative frequency amounts to 7,9 ± 1,56 %. Again, one can use linear interpolation to further refine the interval.*

## *3.3 Confidence intervals in general*

*If the variable is not distributed binomially, one cannot apply the previously mentioned approximation, but has to compute the standard error directly from the individual data. As we estimate averages or totals, we can apply the central limit theorem and determine a α% confidence interval for the estimate h of the characteristic H using this formula:*

$$h \mp u_{1-\alpha/2} \cdot s_h, \qquad (2a)$$

*h* *is the estimate of characteristic H,*
$s_h$ *is the standard error of the estimate h*
$u_{1-\alpha/2}$ *is the quantile of normal distribution.*

***Confidence intervals for average income per capita***

*We collect data about incomes for the whole household. Therefore, the average income per capita is computed as the ratio of 2 random variables, namely y - total of incomes and x - total of persons. Provided simple random sampling without replacement applies and we weight the sample data with weights w, one can determine the α% confidence interval using this formula:*

$$\frac{y_w}{x_w} \pm \frac{u_{1-\alpha/2}}{x_w} \sqrt{\left(1-\frac{n}{N}\right)\frac{n}{n-1}\frac{n}{\sum_{i=1}^{n} w_i}\sum_{i=1}^{n}\left[w_i\left(y_i-\frac{y_w}{x_w}x_i\right)^2\right]} \qquad (2b)$$

$u_{1-\alpha/2}$ *is the quantile of normal distribution (in our case 1,96),*
*n* *the sample size,*
$x_w\,(y_w)$ *are weighted sample totals* $x_w = \frac{n}{\sum w_i}\sum_{i=1}^{n} w_i x_i$ *resp.* $y_w = \frac{n}{\sum w_i}\sum_{i=1}^{n} w_i y_i$

*Although computed confidence intervals in tables III, IV and V were based on this formula assuming simple random sample, the influence of design effect was additionally taken into account. Simplified, it is the influence of complicated sampling scheme on the variability of the estimated characteristic compared to the same result assuming simple random sampling. In reality, as previously described, the sample was stratified at the level of NUTS3 and 4 size-groups of municipalities and was carried out in two stages (see chapter 1.1). Generally the design effect is quantified in compliance with this formula:*

$$deff(h) = s_h^2 / s_h^2\{srs\}, \qquad (3)$$

$s_h^2$ *is the variance of variable h for the real sampling design*

$s_h^2\{srs\}$ *is the variance of variable h for the simple random sample.*

*It is known from the theory that stratification decreases variance, whereas multistage sampling causes estimates with equal observations to be less efficient. Due to the higher total number of dwelling units selected, more CEUs (census enumeration districts) were also (both relatively and absolutely) included. The influence of the above mentioned deff therefore decreased in accordance with the expectation, so stratification and multistage sampling effects were practically balanced, and its values for not rarely occurring income categories for the whole Czech Republic were very close to 1.*

*A modification of formulas (2a) and (2b) was used to compute values in tables III, IV and V. Total variability was in each case decomposed to its components corresponding to each sampling stage.*

*For the values in Table VI, analogous formulas were used which take into account the fact that this is the proportion of frequency estimates depending on the fulfillment of the conditions for inclusion in the calculation of individual indicators.*

*x x x*

*In this publication, it was possible to bring out only some results of the seventh survey on incomes and living conditions in the CR - Living conditions 2020. The data collected in the survey makes it possible to publish various other breakdowns that are not included in this output. For further information contact Information Services (+420 274 052 304 or infoservis@czso.cz).*